

Лекция 5

Разработка плана экспериментальных исследований

Цель лекции: изучить теоретические основы построения плана эксперимента и его применение в научных исследованиях

План лекции:

1. Теоретические основы построения плана эксперимента.
2. Разработка плана эксперимента.
3. Пример проведения полнофакторного экстремального эксперимента

1. Теоретические основы построения плана эксперимента.

Одной из основных задач эксперимента является выявление взаимосвязей между входящими и выходящими параметрами и представление их в виде регрессионной математической модели.

Преимущества использования плана эксперимента:

- 1) резко сокращается число опытов;
- 2) вся схема исследования протекания технологического процесса оказывается значительно формализованной;
- 3) план эксперимента определяет четкую стратегию (последовательность действий), что позволяет принимать обоснованные решения после каждой серии опытов;
- 4) процедура разработки математических моделей значительно упрощается;
- 5) точность математических моделей (их адекватность результатам эксперимента) значительно повышается;
- 6) разработаны математические модели позволяют глубже выявить механизм явления и определить оптимальные значения сразу всех факторов x_i (потому что они действуют на реальный процесс и изменяются одновременно), при которых значение всех функций оптимизации, также оптимальны.

План эксперимента – совокупность данных определяющих число, условия и порядок проведения опытов.

Планирование эксперимента – выбор плана эксперимента, удовлетворяющего заданным требованиям, совокупность действий направленных на разработку стратегии экспериментирования (от получения априорной информации до получения работоспособной математической модели или определения оптимальных условий).

Цель планирования эксперимента – нахождение таких условий и правил проведения опытов при которых удастся получить надежную и достоверную информацию об объекте с наименьшей затратой труда, а также представить эту информацию в компактной и удобной форме с количественной оценкой точности.

Планирование эксперимента включает следующие составляющие элементы:

- 1) определение необходимого числа опытов или наблюдений;
- 2) выбор забора плана эксперимента;
- 3) разработку плана эксперимента (математическая модель и матрица планирования).

Пусть интересующее нас свойство (Y) объекта зависит от нескольких (n) независимых переменных (X_1, X_2, \dots, X_n) и мы хотим выяснить характер этой зависимости – $Y=F(X_1, X_2, \dots, X_n)$, о которой мы имеем лишь общее представление.

Величина Y – называется «отклик», а сама зависимость $Y=F(X_1, X_2, \dots, X_n)$ – "функция отклика".

Отклик должен быть определен количественно. Однако могут встречаться и качественные признаки Y . В этом случае возможно применение рангового подхода.

Независимые переменные X_1, X_2, \dots, X_n – иначе факторы, также должны иметь количественную оценку.

Область значений факторов X , в которой находятся точки, отвечающие условиям проведения опытов используемого плана эксперимента, называется областью планирования.

Чаще всего область планирования задаётся интервалами возможного изменения факторов

$$x_{i\min} \leq x_i \leq x_{i\max}, \quad i=1, 2, \dots, n$$

Планирование эксперимента



Параметр оптимизации – обобщенный параметр представляемый как функция от некоторого количества входящих параметров.

Фактор - независимая переменная величина, принимающая в некоторый момент времени определенное значение, влияет на параметр оптимизации.

2. Разработка плана эксперимента.

Полным факторным экспериментом (ПФЭ)

называется такой эксперимент, при реализации которого определяется значение функции цели для всех возможных сочетаний варьирования (изменения) факторов.

Общее число различных комбинаций уровней в ПФЭ для k факторов можно вычислить как $n = q_1 \cdot q_2 \cdot \dots \cdot q_k$, где q_i – число уровней i -го фактора.

Если имеется k факторов, каждый из которых может устанавливаться на q уровнях, то для осуществления полного факторного эксперимента необходимо поставить $n = q^k$ опытов.

Общее уравнение для получения математической модели объекта при планах многофакторного анализа имеет вид

$$y = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n \sum_{j=2}^n b_{ij} x_i x_j + \dots$$

где b_i – оценка генеральных параметров модели.

Для двухфакторного анализа уравнение примет вид

$$y = b_0 + b_1 x_1 + b_2 x_2$$

Планирование, проведение и обработка результатов ПФЭ состоят из следующих обязательных этапов:

- 1) кодирование факторов;
- 2) составление плана матрицы эксперимента;
- 3) реализация плана эксперимента;
- 4) проверка воспроизводимости опытов;
- 5) оценка значимости коэффициентов регрессии;
- 6) проверка адекватности линейной модели.

1) кодирование факторов

Кодирование факторов необходимо для перевода натуральных факторов в безразмерные величины, чтобы иметь возможность построить стандартную ортогональную план-матрицу эксперимента.

После выбора факторов для каждого из них устанавливается основной уровень (т.е. исходное значение) и интервал варьирования (изменения). Прибавление интервала варьирования к основному уровню даёт *верхний*, а вычитание – *нижний* уровень фактора. Удобно, чтобы верхний уровень соответствовал **+1**, нижний **-1**, а основной **0**.

Таблица – Уровни варьирования факторов

Параметр	Минимальное значение	Среднее значение	Максимальное значение
	-1	0	+1
	-1	0	+1

2) составление плана матрицы эксперимента

Таблица – План матрица полнофакторного эксперимента

Серия опытов	Уровни варьирования		
	X1	X2	X3
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	+	+	+
6	+	-	+
7	-	+	+
8	-	-	+

3) реализация плана эксперимента

Таблица – Фрагмент конечных результатов проведения эксперимента

Серия	Y1	Y2	Y3	...	Y19	Y 20	Среднее
1	y ₁₁	y ₁₂	y ₁₃	...	y ₁₁₉	y ₁₂₀	y1
2	y ₂₁	y ₂₂	y ₂₃	...	y ₂₁₉	y ₂₂₀	y2
...
7	y ₇₁	y ₇₂	y ₇₃	...	y ₇₁₉	y ₇₂₀	y7
8	y ₈₁	y ₈₂	y ₈₃	...	y ₈₁₉	y ₂₂₀	y8

20 (n_{∂}) – количество наблюдений по сериям, определяется по формуле

$$n_{\partial} = \frac{t_{\beta}^2 \cdot \sigma^2}{\varepsilon^2},$$

Таблица – Результаты наблюдений эксперимента

Номер наблюдения	Уровни фактора						
	1	2	...	j	...	k	
1	y_{11}	y_{12}	...	y_{1j}	...	y_{1k}	
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}	
Суммы $Y_j = \sum_{i=1}^{n_j} y_{ij}$	Y_{i1}	Y_{i2}	...	Y_{ij}	...	Y_{ik}	$Y = \sum_{j=1}^k Y_{ij}$
Число наблюдений	n_1	n_2	...	n_j	...	n_k	$N = \sum_{j=1}^k n_j$
Средние $\bar{y}_{ij} = Y_j/n_j$	\bar{y}_{i1}	\bar{y}_{i2}	...	\bar{y}_{ij}	...	\bar{y}_{ik}	$\bar{\bar{y}} = Y/N$
Квадраты сумм	Y_{i1}^2	Y_{i2}^2	...	Y_{ij}^2	...	Y_{ik}^2	Y^2

4) проверка воспроизводимости опытов;

Воспроизводимость состоит из
следующих этапов:

1) оценка дисперсий. Дисперсии для
каждой серии опытов определяется по
следующей зависимости:

$$s_i^2 = \frac{1}{m} \cdot \sum_{j=1}^m (y_{ij} - \bar{y}_i)^2$$

где m – количество наблюдений в серии.

2) Определение воспроизводимости по критериям:

- критерий Фишера (при не одинаковом числе опытов) :

$$F = \frac{S_A^2}{S_{от}^2} \begin{array}{l} \leftarrow \text{Дисперсия между выборками} \\ \leftarrow \text{Дисперсия внутри выборок} \end{array}$$

По таблице выбираем значение по уровню значимости и числу степеней свободы и сравниваем с расчетным:

$$F_{расч} < F_{табл}$$

Если условие соблюдается, то эксперимент воспроизводимый.

Значения критерия Фишера (F-критерия) для уровня значимости $\alpha = 5\%$

ν_1 – число степеней свободы большей дисперсии; ν_2 – число степеней свободы меньшей дисперсии

ν_2	ν_1											
	1	2	3	4	5	6	7	8	9	10	15	20
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	245,9	248,0
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43	19,45
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12

- критерий Кохрена (при одинаковом
числе опытов)

$$G_{расч} = \frac{\max S_i^2}{\sum_1^r S_i^2}$$

Критические значения коэффициента Кохрена (G-критерия) для доверительной вероятности $p = 95\%$ и числе степеней свободы ν

Число измерений, k	Число степеней свободы, ν										
	1	2	3	4	5	6	8	10	16	36	∞
2	9985	9750	9392	9057	8772	8534	8159	7880	7341	6602	5000
3	9669	8709	0797	7454	7071	6771	6333	6025	5466	4748	3333
4	9065	7679	6841	6287	5895	5598	5175	4884	4366	3720	2500
5	8412	6838	5981	5441	5065	4783	4387	4118	3645	3066	2000
6	7808	6161	5321	4803	4447	4184	3817	3568	3135	2612	1667
7	7271	5612	4800	4307	3974	3726	3384	3154	2756	2278	1429
8	6798	5157	4377	3910	3595	3362	3043	2829	2462	2022	1250
9	6385	4775	4027	3584	7276	3067	2768	2568	2226	1820	1111
10	6020	4450	3733	3311	3029	2823	2541	2353	2032	1655	1000
12	5410	3924	3264	2880	2624	2439	2187	2020	1737	1403	0833
15	4709	3346	2758	2419	2195	2034	1815	1671	1429	1144	0667
20	3894	2705	2205	1921	1735	1602	1422	1303	1108	0879	0500
24	3434	2354	1907	1656	1493	1374	1216	1113	0942	0743	0417
30	2929	1980	1593	1377	1237	1137	1001	0921	0771	0604	0333
40	2370	1576	1259	1082	0968	0887	5950	0713	0595	0462	0250
60	1737	1131	0895	0765	0682	0623	0552	0497	0411	0316	0167
120	0998	0632	0495	0419	0371	0337	0292	0266	0218	0165	0083
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

Все значения G-критерия меньше единицы, поэтому в таблице приведены лишь десятичные знаки, следующие после запятой, перед которой при пользовании таблицей нужно ставить ноль целых.

Например, при $k = 6, \nu = 3$ имеем $G_{0,95} = 0,5321$.

3. Пример проведения полнофакторного экстремального эксперимента

Название параметра	Минимальное значение	Среднее значение	Максимальное значение	Закон распределения
Параметры внешней среды				
Объем отправки, т	1	5	10	нормальный
Расстояние перевозки, км	50	150	300	нормальный
Управляемые параметры				
Техническая скорость, км/ч	20	23	26	нормальный
Коэффициент использования грузоподъемности	0,5	0,75	0,99	нормальный

Параметр оценки – время доставки (стремится к минимуму)